

Revealing the specificity of human H1 influenza A viruses to complex N-glycans

Angeles Canales^{a*}, Javier Sastre^b, Jose M. Orduña^c, Cindy M. Spruit^d, Javier Pérez-Castells^c, Gema Domínguez^c, Kim M. Bouwman^d, Roosmarijn van der Woude^d, Francisco Javier Cañada^{b,e}, Corwin M. Nycholat^f, James C. Paulson^f, Geert-Jan Boons^{d,g}, Jesús Jiménez-Barbero^{e,h,i,j}, Robert P. de Vries^{d*}.

a Department of Organic Chemistry, Faculty of Chemistry, Universidad Complutense de Madrid, Avd. Complutense s/n, 28040, Madrid, Spain

b Centro de Investigaciones Biológicas Margarita Salas, CSIC, C/Ramiro de Maetzu 9, 28040, Madrid, Spain

c Department of Chemistry and Biochemistry Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities Urbanización Montepríncipe, 28660 Boadilla del Monte, Madrid, Spain

d Department of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, 3584 CG The Netherlands

e Centro de Investigación Biomédica En Red de Enfermedades Respiratorias, Madrid, Spain.

f Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA, 92037;

g Complex Carbohydrate Research Center, University of Georgia, Athens, USA.

h CIC bioGUNE, Bizkaia Science and Technology Park, 48160, Bilbao, Spain

i IKERBASQUE, Basque Foundation for Science, 48009, Bilbao, Spain

j Department of Organic Chemistry, II Faculty of Science and Technology University of the Basque Country, EHU-UPV, 48940 Leioa, Spain.

CONTENTS:

1. Mirage glycan array report.
2. Figure S1. Supporting tissue binding experiments.
3. Synthesis of trisaccharide probe **19**, synthetic procedures and characterization data
4. Figure S2. TriLacNAc biantennary N-glycan **13**, ^1H - ^{13}C -HSQC NMR spectrum
5. Figure S3. DiLacNAc biantennary N-glycan **10**, ^1H - ^{13}C -HSQC NMR spectrum
6. Figure S4. Expansions of the STD spectra of the TriLacNAc biantennary N-glycan acquired in the presence of H6 G225D, H1 A/NC/20/99, H1 A/Bris/18 and H1 A/CA/04/09 hemagglutinins.
7. Figure S5. Proton signals that are identified in STD spectrum as interaction points with A/CA/04/09.
8. Figure S6. Comparison of the STD spectra of the TriLacNAc and diLacNAc biantennary glycans when bound to H1 A/CA04/09 and H6N1 G225D mutant.
9. Figure S7. Model of the bidentate binding of a triLacNAc glycan recognizing two sites of the same hemagglutinin trimer.
10. Figure S8. Comparison of the STD spectra (acetyl group region) of the TriLacNAc biantennary N-glycan in the presence of the four hemagglutinins characterized in this work.
11. Figure S9. Control STD of the probe 19 acquired in the absence of lanthanide to characterize the interaction with H1 A/CA/04/09

Table S1

Supplementary Glycan Microarray Information, Based on MIRAGE Guidelines
(doi:10.3762/mirage.3)

	Description*
1. Sample: Glycan Binding Sample	
Description of Sample	<p>HAs were expressed and assayed as previously described (Nemanichvili et al., 2018).</p> <p>Fig 1B A/Duck/UKR/63 A/NC/20/99 Fig 1C A/Cal/04/09 (Cal/04/09) A/Bris/18 A/Taiwan/1/13 H6N1 G225D</p> <p>HA-encoding cDNA was cloned into the pCD5 expression vector as described previously. The resulting expression vector encodes an HA protein containing a heterologous signal peptide, a C-terminal trimerization domain (GCN4), sfGFP, and a double strep-tag, and lacking the transmembrane and cytoplasmic domains.</p> <p>The HA proteins were expressed in HEK293S GnT1(-) cells, purified from the cell culture supernatants using streptactin beads, and quantified as described previously. (Nemanichvili et al., 2018)</p>

Assay protocol	HA (50 µg/mL) were precomplexed with anti-streptag and goat anti-mouse-555 in 50 µL for 15 min on ice and incubated on the array for 90 min at room temperature.
2. Glycan Library	
Glycan description for defined glycans	Broszeit et al., 2021
3. Printing Surface; e.g., Microarray Slide	
Description of surface	NHS-activated glass (NEXTERION® Slide H,)
Manufacturer	Schott Inc.
Covalent Immobilization	NHS-ester for amine functional groups (reacting with terminal Asn)
4. Arrayer (Printer)	
Description of Arrayer	Non-contact microarray printer, sciFLEXARRAYER S3, Scienion Inc.
Dispensing mechanism	Non-contact, one nozzle
Glycan deposition	400 pL of 100 µM, 40 fmol, 6 replicates
Printing conditions	100 µM printing concentration, 20 °C, 50% humidity, blocking with 5 mM ethanolamine in TRIS buffer (pH 9, 100 mM) for 1 h at 50 °C. Slides were rinsed with DI water.
5. Glycan Microarray with “Map”	
Array layout	24 subarrays (3 x 8) were printed per slide. See Figure Sx for layout of the subarray.
Glycan identification and quality control	For identification and quality control see Broszeit et al., 2021
6. Detector and Data Processing	
Scanning hardware	Innopsys Innoscan 7200
Scanner settings	Iterative scans at 532 nm, laserpower at 1%, 5%, 10%, 50% and 100% (532 nm). The lower laserpowers were used to avoid overexposure.
Image analysis software	GenePix Pro 7 software

Table S3. Compound map of the array. Locations of compound 1 to 15 are indicated (compound numbers refer to Figure 1A). A 532, and x indicate printing of Atto 532, and compounds unrelated to this project, respectively.

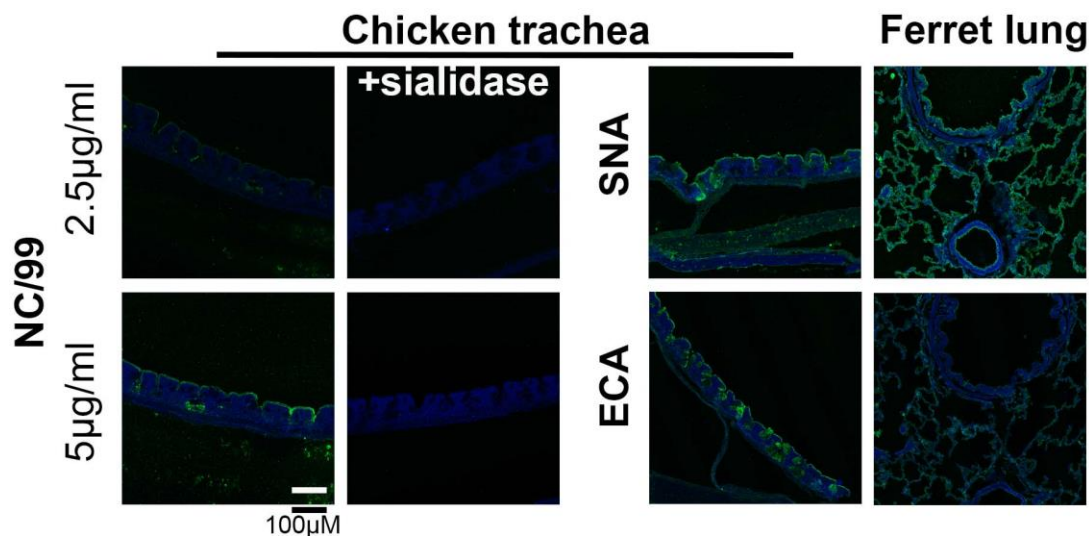


Figure S1. Concentration and sialic acid dependent binding of NC/20/99 to the chicken trachea. Plant lectins SNA and ECA staining of the chicken trachea and the ferret lung

2. Synthesis of trisaccharide probe 19

All chemicals were obtained from Aldrich/Merck (St. Louis, MO, USA) and VWR (Radnor, PA, USA), Fluorochem (Derbyshire, UK). TLC analyses were performed on Merck silica gel 60 F254 plates using phosphomolybdic acid or anisaldehyde and heat for detection. Silica gel NORMASIL 60 40–63 μm was used for flash chromatography. NMR spectra were recorded on a Bruker spectrometer (400 MHz or 300 MHz for ^1H and 100 MHz or 75 MHz for ^{13}C), (Billerica, MA, USA). Chemical shifts are reported in δ ppm referenced to CDCl_3 ($\delta = 7.26$ for ^1H and 77.00 for ^{13}C), CD_3OD ($\delta = 3.31$ for ^1H and 49.00 for ^{13}C), or D_2O ($\delta = 4.79$ for ^1H). Bidimensional spectra (HMQC, HMBC, COSY, and NOESY) were recorded in order to carry out the assignment. Infrared spectra were done in a Perkin-Elmer spectrum 100 (Agilent, Santa Clara, CA, USA). Specific optical rotation was measured in a Polarimeter Anton Paar MCP 100 (Anton Paar, Graz, Austria). Melting points of solid compounds were determined using a Stuart Scientific Melting Point Apparatus SMP3 (Stuart, Staffordshire, UK). Microanalyses were done on a LECO CHNS-932 (LECO, St. Joseph, MI, USA). Absorbance of *p*-nitrophenoxide released in the enzymatic reactions was measured at 405 nm in a Perkin-Elmer Lambda25 (PerkinElmer, Waltham, MA)

Tetraethyl 2,2',2'',2'''-(((4-iodopyridine-2,6-diyl)bis(methylene))bis(azanetriyl))tetraacetate, 16. It was prepared following the procedure described in the literature.^[1] The spectroscopic data were in accordance with those previously reported. ^1H RMN (400 MHz, CDCl_3) δ (ppm): 7.91 (s,

2H, 2xPy-H), 4.17 (q, $J = 7.1$ Hz, 8H, 4xOCH₂), 3.98 (s, 4H, 2xPyCH₂N), 3.59 (s, 8H, 4xNCH₂CO), 1.28 (t, $J = 7.1$ Hz, 12H, 4xCH₃).

Tetraethyl 2,2',2'',2'''-(((4-((trimethylsilyl)ethynyl) pyridine-2,6-diyl) bis(methylene)) bis(azanetriyl))tetraacetate. In a glass pressure tube were placed 1.012 g (1.67 mmol) of tetraethyl 2,2',2'',2'''-(((4-iodopyridine-2,6-diyl)bis(methylene))bis(azanetriyl))tetraacetate (**16**) 1.2 mg (6.7 μ mol) of copper iodide, 23.4 mg (0.033 mmol) of PdCl₂(PPh₃)₂, and 10 mL of triethylamine. This mixture was degassed with argon and later, 0.31 mL (2 mmol) of ethynyltrimethylsilane were added. The reaction was warmed up to 40 °C and stirred for 4 hours. After that time, the mixture was filtered through a Celite™ pad and a short chromatography was carried out in silica gel using a mixture of hexane and EtOAc (3:2) as mobile phase, affording 817 mg of the product (85% yield) as a yellow oil. ¹H RMN (400 MHz, CDCl₃) δ (ppm): 7.49 (s, 2H, 2xPy-H), 4.15 (q, $J = 7.1$ Hz, 8H, 4xOCH₂), 3.99 (s, 4H, 2xPyCH₂N), 3.58 (s, 8H, 4xNCH₂CO), 1.26 (t, $J = 7.1$ Hz, 12H, 4xCH₃), 0.23 (s, 9H, (CH₃)₃Si). ¹³C RMN (100 MHz, CDCl₃) δ (ppm): 171.4 (4xCO), 158.9 (2xC-Py), 132.7 (C Py), 123.7 (2xCH-Py), 103.0 (PyC \equiv), 99.4 (\equiv C-TMS), 60.8 (4xCH₂O), 60.0 (2xCH₂Py), 55.2 (4xCH₂CO), 14.5 (4xCH₃), -0.004 (3xCH₃Si). IR (NaCl, cm⁻¹): 2976, 2165, 1746, 1594, 1190. Anal. Calcd. for C₂₈H₄₃N₃O₈Si: C, 58.21; H, 7.50; N, 7.27. Found: C, 58.32; H, 7.66; N, 7.39.

Tetraethyl 2,2',2'',2'''-(((4-ethynylpyridine-2,6-diyl)bis(methylene))bis(azanetriyl))tetraacetate, 17. 0.22 mmol (127 mg) of the previous compound were dissolved in 1.25 mL of THF, and 0.24 mL of 1 M TBAF (tetrabutylammonium fluoride) solution in THF were added. After stirring 1 hour at rt, no starting material was detected by NMR analysis. Subsequently, the solvent was removed by evaporation at low pressure and a mixture of water (10 mL), and EtOAc was added (10 mL). The aqueous phase was extracted with EtOAc (3x10 mL), dried first with 20 mL of brine solution and, later over anhydrous sodium sulphate. The solvent was removed by evaporation at low pressure, obtaining 78 mg the desired product **17** as a yellow oil (70% yield). ¹H RMN (400 MHz, CDCl₃) δ (ppm): 7.57 (s, 2H, 2xPy-H), 4.16 (q, $J = 7.1$ Hz, 8H, 4xOCH₂), 4.02 (s, 4H, 2xPyCH₂N), 3.60 (s, 8H, 4xNCH₂O), 3.23 (s, 1H, H-C \equiv), 1.27 (t, $J = 7.1$, 12H, 4xCH₃). ¹³C RMN (100 MHz, CDCl₃) δ (ppm): 171.1 (4xCO), 158.8 (2xC-Py), 131.5 (C-Py), 123.7 (2xCH-Py), 81.3 (PyC \equiv), 77.1 (\equiv C-H), 60.6 (4xCH₂O), 59.6 (2xCH₂Py), 54.9 (4xCH₂CO), 14.2 (4xCH₃). IR (NaCl, cm⁻¹): 3260, 2981, 2110, 1744, 1214. Anal. Calcd. for C₂₅H₃₅N₃O₈: C, 59.39; H, 6.98; N, 8.31. Found: C, 59.21; H, 6.87; N, 8.28.

N-Acetyl- α -neuraminosyl-(2-6)- β -D-galactopyranosyl-(1-4)- β -D-glucopyranosyl azide, 18. Following a described procedure,^[2] 6-sialyl lactose (500 mg, 0.789 mmol), sodium azide (8 mmol, 513 mg) and DIPEA (8 mmol, 1.4 mL) were dissolved in 3.15 mL of MQ water and placed in a 0

°C ice bath. Then, 2.567 mmol (434 mg) of 2-chloro-1,3-dimethyl-1-*H*-imidazolium chloride were added, and reaction was kept at 0 °C for 1 hour. Water was removed by co-evaporation with acetonitrile at low pressure. Then 5 mL of DMF were added, and the solid was resuspended using ultrasounds and filtered. DMF was removed by co-evaporation with toluene at low pressure. After the elimination of the solvent, 5 mL of MQ water and 5 mL of DCM were added. The aqueous phase was extracted five times with DCM (5 mL, each), and the aqueous phase was evaporated at low pressure affording a grey solid which was dissolved in MeOH (2 mL). Then, 10 mL of isopropanol were added, and the mixture was stirred for 1 hour at rt. The resulting white solid precipitate was filtered and washed with isopropanol, acetone and hexane, affording 500 mg of 6'-sialyl lactose-1-azide (*N*-acetyl- α -neuraminosyl-(2-6)- β -D-galactopyranosyl-(1-4)- β -D-glucopyranosyl azide, **3**), 96% yield. ¹H NMR (400 MHz, D₂O) δ (ppm): 4.66 (d, *J* = 8.8 Hz, 1H, H-1'), 4.32 (d, *J* = 7.8 Hz, 1H, H-1), 3.92 – 3.69 (m, 8H), 3.62 – 3.42 (m, 10H), 3.33 (t, *J* = 8.8 Hz, 1H), 2.60 (dd, *J* = 12.4, 4.5 Hz, 1H, H-3''), 1.92 (s, 3H, CH₃), 1.62 (t, *J* = 12.2 Hz, 1H, H-3''). ¹³C NMR (100 MHz, D₂O) δ (ppm): 174.8 (CO), 173.4 (CO), 103.1 (CH-1'), 100.2 (C-2''), 89.7 (CH-1), 78.9 (CH), 76.5 (CH), 74.5 (CH), 73.7 (CH), 72.5 (CH), 72.3 (2xCH), 71.8 (CH), 70.5 (CH), 68.5 (CH), 68.3 (2xCH), 63.6 (CH₂), 62.6 (CH₂), 59.9 (CH₂), 51.7 (CH-5''), 40.0 (CH₂-C3''), 22.0 (CH₃). IR (NaCl, cm⁻¹): 3280, 2135, 1750, 1621, 1240. Anal. Calcd. for C₂₃H₃₈N₄O₁₈ Calc: C, 41.95; H, 5.82; N, 8.51. Found: C, 42.19; H, 5.87; N, 8.66. ²⁵[α]_D: -6.54 (32.5 mg/mL, methanol). Mp. 193 °C dec.

Tetraethyl 2,2',2'',2'''-(((4-(1-(*N*-acetyl- α -neuraminosyl-(2-6)- β -D-galactopyranosyl-(1-4)- β -D-glucopyranosyl)-1*H*-1,2,3-triazol-4-yl)pyridine-2,6-diyl)bis(methylene))bis(azanetriyl))tetraacetate. 0.228 mmol (150 mg) of *N*-acetyl- α -neuraminosyl-(2-6)- β -D-galactopyranosyl-(1-4)- β -D-glucopyranosyl azide (**18**) dissolved in 1.2 mL of argon degassed methanol, were added in an argon degassed mixture of THF (1.2 mL) which contains 0.251 mmol (127 mg) of the starting alkyne (**17**) and 0.456 mmol (90 mg) of sodium ascorbate dissolved in 0.6 mL of water. Finally, 0.046 mmol (11 mg) of copper sulphate dissolved in 0.6 mL of degassed water were added. The reaction was kept overnight with stirring, under argon atmosphere and in dark. Later, the reaction mixture was evaporated and purified through a C18 silica gel column using acetonitrile, water and TFA (0.1 %) as mobile phase in a gradient of acetonitrile (5 to 35%) Fractions with product were lyophilized affording 241 mg of the desired product as a pale blue solid (91% yield). ¹H NMR (400 MHz, D₂O) δ (ppm): δ 8.94 (s, 1H, Triazole-H), 8.13 (s, 2H, 2xPy-H), 5.83 (d, *J* = 9.2 Hz, 1H, H-1), 4.42-4.40 (m, 5H, H-1' and 2xPyCH₂N), 4.08 (q, *J* = 7.1 Hz, 8H, 4xOCH₂), 3.97 – 3.44 (m, 27H), 2.63 (dd, *J* = 12.5, 4.2 Hz, 1H, H-3''), 1.91 (s, 3H, CH₃CO), 1.71 (t, *J* = 12.1 Hz, 1H, H-3''), 1.12 (t, *J* = 7.1 Hz, 12H, 4xCH₃). ¹³C NMR (100 MHz, D₂O) δ (ppm): 174.8 (CO), 172.6 (4xCO), 171.9 (CO), 153.8 (2xC-Py), 145.7 (C-Py or C-Triazole), 142.8 (C-Py or C-Triazole), 125.7 (CH-Triazole),

119.9 (2xCH-Py), 103.1 (CH-1'), 99.3 (C-2''), 87.4 (CH), 78.6 (CH), 77.6 (CH), 74.6 (CH), 73.6 (CH), 72.6 (CH), 72.3 (CH), 71.9 (CH), 71.2 (CH), 70.6 (CH), 68.4 (CH), 68.3 (CH), 67.9 (CH), 63.6 (CH₂), 62.8 (CH₂), 62.1 (4xCH₂O), 59.7 (CH₂), 55.7 (4xCH₂CO), 55.5 (2xCH₂Py), 51.6 (CH-5''), 39.4 (CH₂-3''), 22.0 (CH₃CO), 13.2 (4xCH₃). IR (NaCl, cm⁻¹): 3284, 3100, 1749, 1618, 1238. Anal. Calcd. for C₄₈H₇₃N₇O₂₆: C, 49.52; H, 6.32; N, 8.42. Found: C, 49.61; H, 6.39; N, 8.57. ²⁵[α]_D: -14.6 (23.3 mg/mL, methanol). Mp. 141 – 143 °C.

2,2',2'',2'''-(((4-(1-(N-acetyl-α-neuraminosyl-(2-6)-β-D-galactopyranosyl-(1-4)-β-D-glucopyranosyl)-1H-1,2,3-triazol-4-yl)pyridine-2,6-diyl)bis(methylene))bis(azanetriyl))tetraacetic acid, 19. 0.043 mmol (50 mg) of the previous compound were dissolved in 1 mL of MeOH and, subsequently, 285 μL of NaOH 3.3 M (0.946 mmol) in MeOH were added. This mixture was stirred for 1 hour at rt. Then, 1 mL of MQ water was added and, after stirring 30 minutes, all solvents were removed by evaporation at low pressure. The product was purified through a reverse column (C18) using a mixture of MQ water and acetonitrile 95:5. All fractions containing product were lyophilized affording 38 mg of **19** as a pale blue solid (77% yield). ¹H NMR (400 MHz, D₂O) δ (ppm): 8.70 (s, 1H, Triazole-H), 7.74 (s, 2H, 2xPy-H), 5.79 (d, *J* = 9.0 Hz, 1H, H-1), 4.43 (d, *J* = 7.8 Hz, 1H, H-1'), 4.15 – 3.46 (m, 23H), 3.19 (s, *J* = 10.1 Hz, 8H), 2.65 (dd, *J* = 12.2, 4.3 Hz, 1H, H-3''), 1.93 (s, 3H, CH₃), 1.68 (t, *J* = 12.2 Hz, 1H, H-3'''). ¹³C NMR (100 MHz, D₂O) δ (ppm): 179.6 (CO), 174.9 (4xCO), 173.5 (CO), 156.0 (2xC-Py), 145.4 (C-Py or C-Triazole), 138.8 (C-Py or C-Triazole), 123.5 (CH-Triazole), 119.1 (2xCH-Py), 103.2 (CH-1'), 100.3 (C-2''), 87.3 (CH-1), 78.7 (CH-4), 77.6 (CH-5), 74.7 (CH-3), 73.7 (CH-5'), 72.5 (CH-3'), 72.3 (CH-6''), 71.9 (CH-2, CH-8''), 70.8 (CH-2'), 68.4 (CH-4', CH-4'', CH-7''), 63.5 (CH₂-6'), 62.6 (CH₂-9''), 59.9 (CH₂-6), 59.1 (2xCH₂Py), 58.1 (4xCH₂CO), 51.8 (CH-5'), 40.1 (CH₂-3''), 22.0 (CH₃). IR (NaCl, cm⁻¹): 3291, 3092, 1740, 1612, 1500. Anal. Calcd. for C₄₀H₅₇N₇O₂₆: C, 45.67; H, 5.46; N, 9.32. Found: C, 45.41; H, 5.43; N, 9.29. ²⁵[α]_D: -4.14 (31.7 mg/mL, D₂O). Mp. 190 – 192 °C, dec.

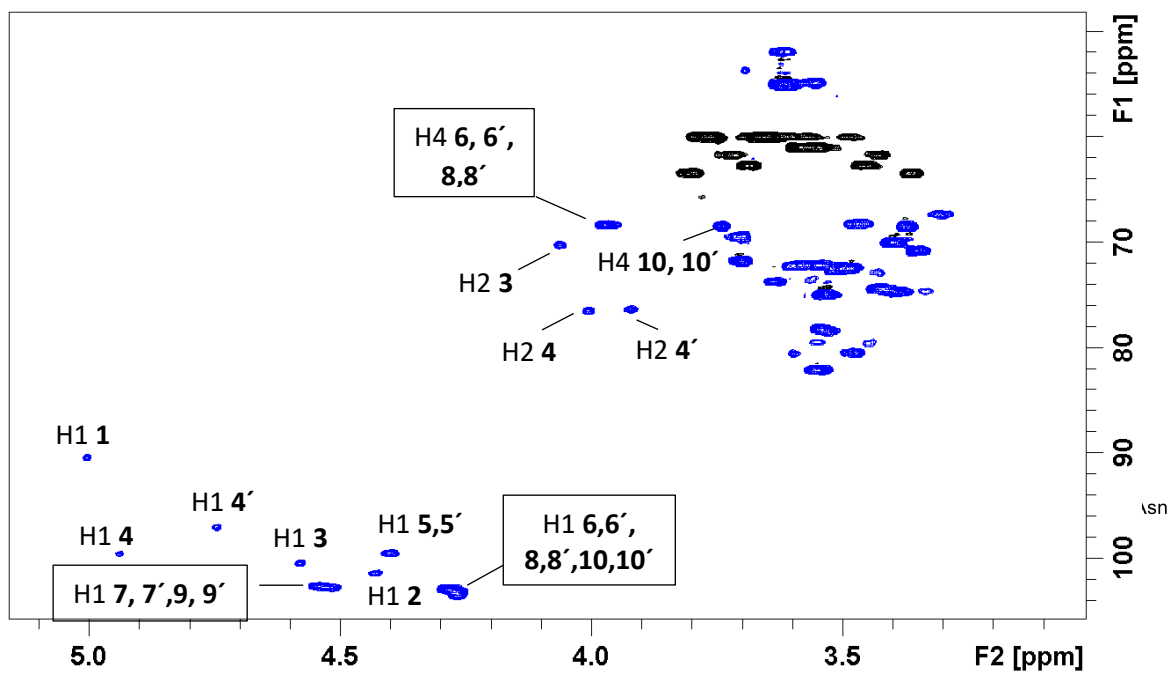
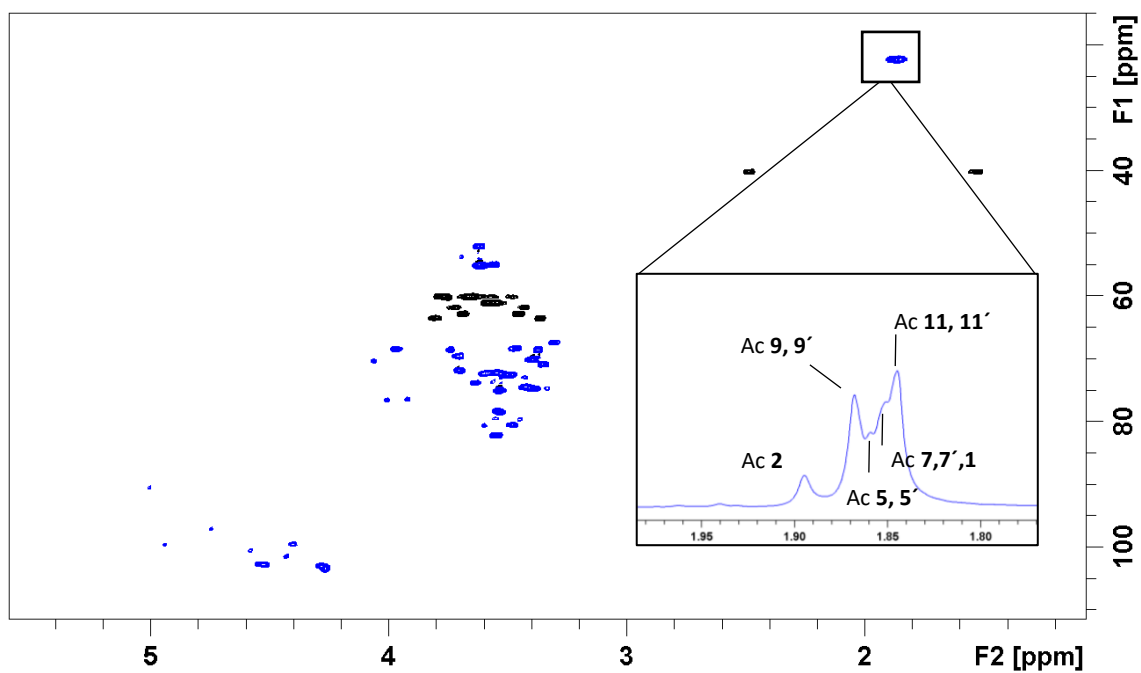
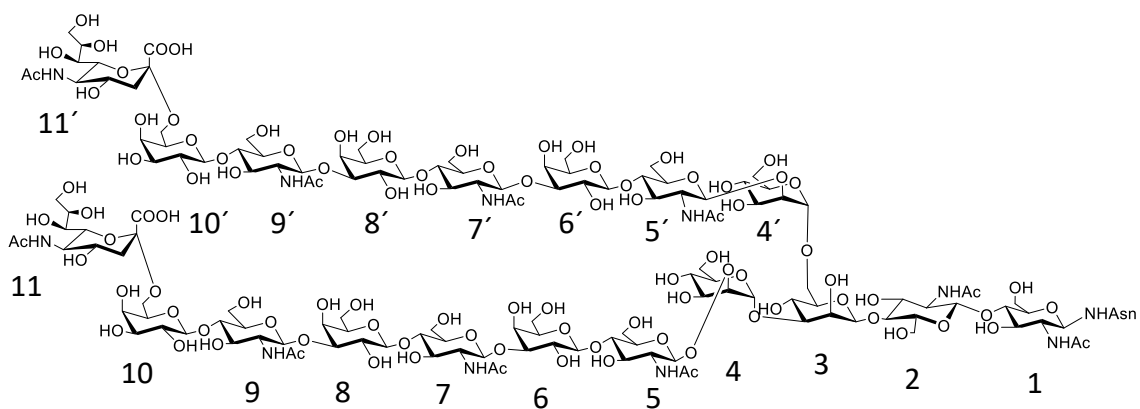


Figure S2. TriLacNAc biantennary N-glycan 13, ^1H - ^{13}C -HSQC NMR spectrum

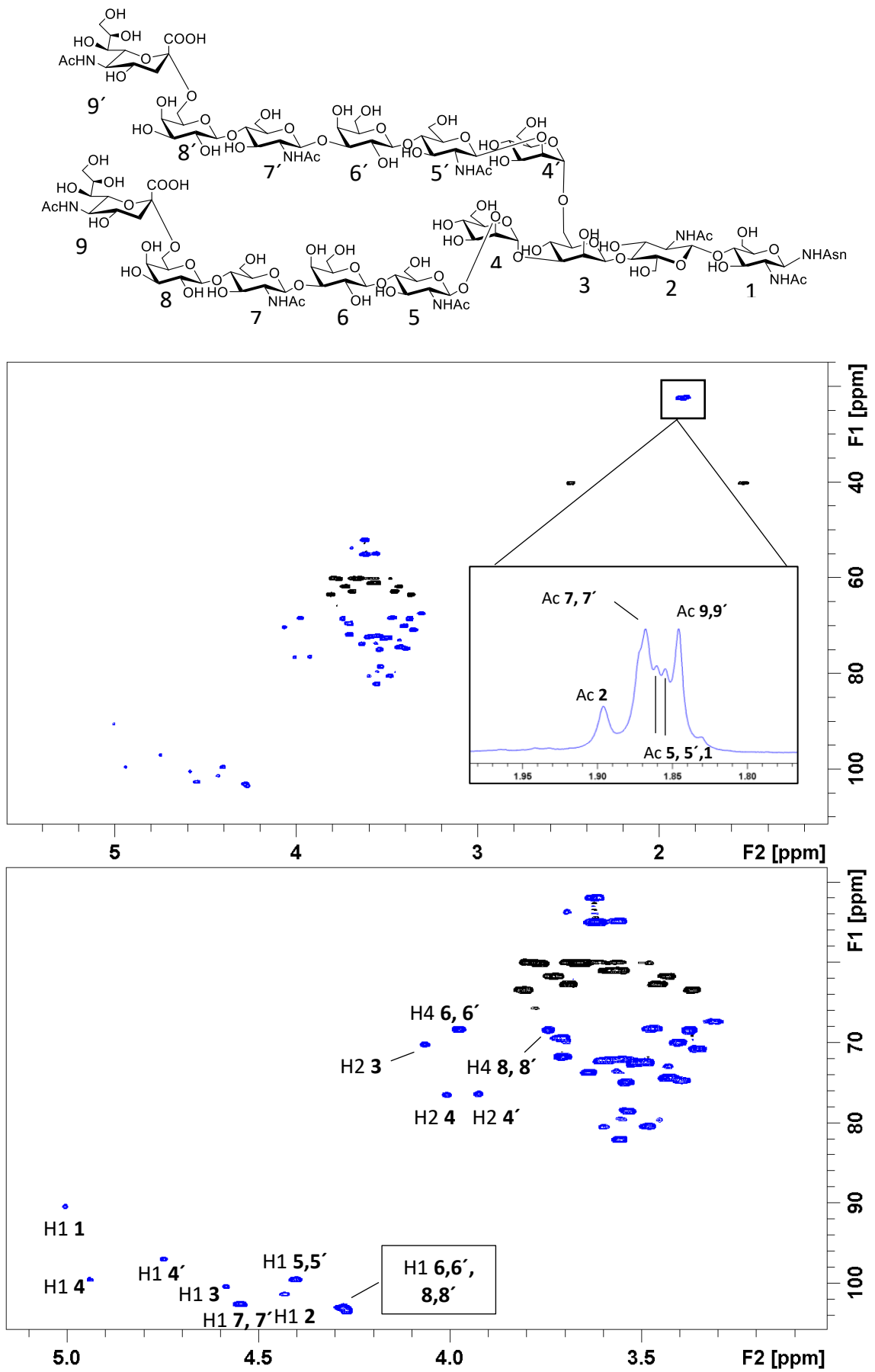


Figure S3. DiLacNAc biantennary N-glycan 10, ^1H - ^{13}C -HSQC NMR spectrum

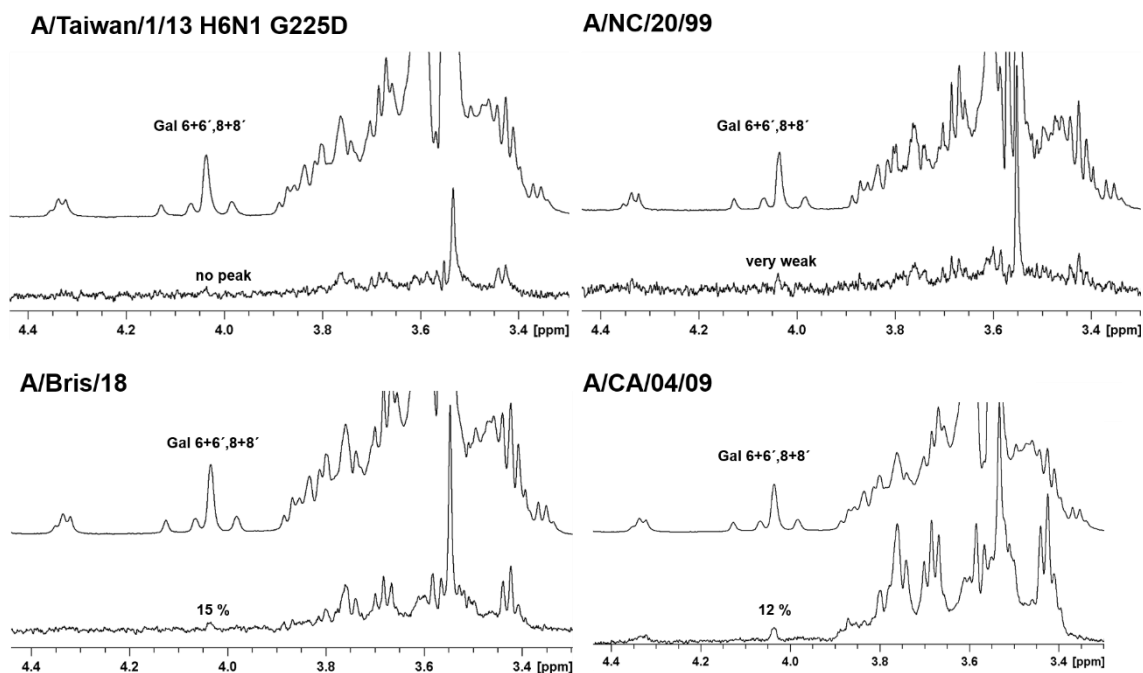


Figure S4. Expansions of the STD spectra of the TriLacNAc biantennary N-glycan acquired in the presence of H6 G225D, H1 A/NC/20/99, H1 A/Bris/18 and H1 A/CA/04/09 hemagglutinins. The STD effects are normalized to the highest signal that is the acetyl group of the sialic acid.

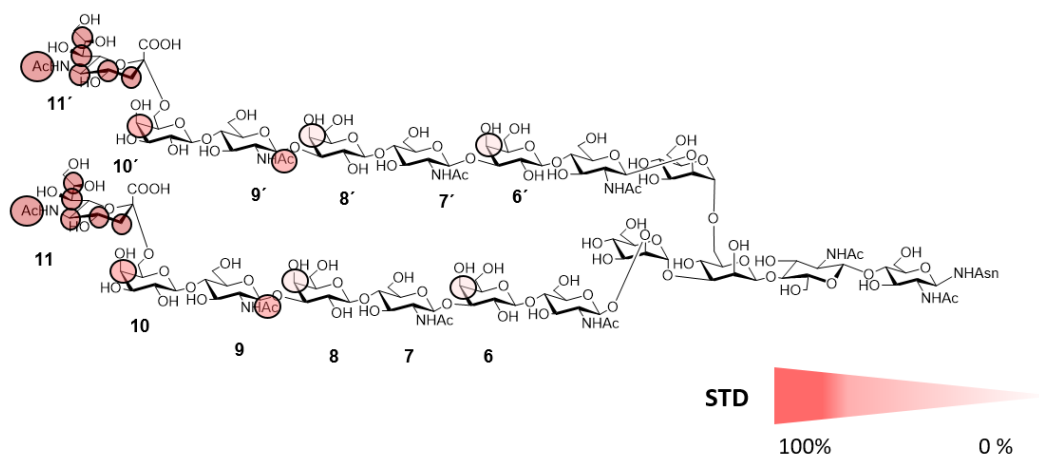


Figure S5. Proton signals that are identified in STD spectrum as interaction points with A/CA/04/09. Signals of internal galactoses 6,6',8,8' overlapped and therefore the contribution of 6,6' and 8,8' to the experimental STD can not be properly quantified.

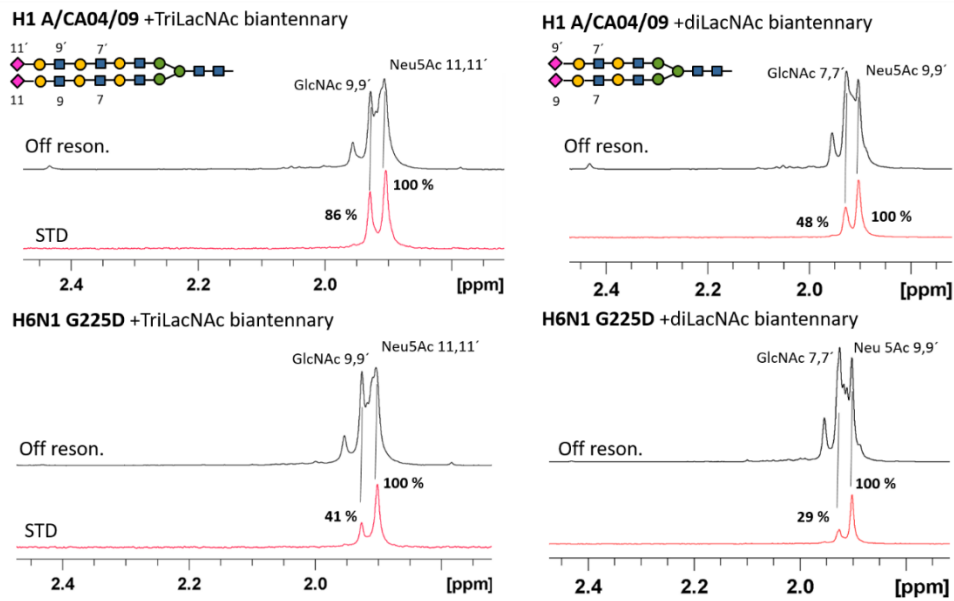


Figure S6. Comparison of the STD spectra of the TriLacNac and diLacNac biantennary glycans when bound to H1 A/CA04/09 and H6N1 G225D mutant. The STD effects are normalized to the highest signal that is the acetyl group of the sialic acid.

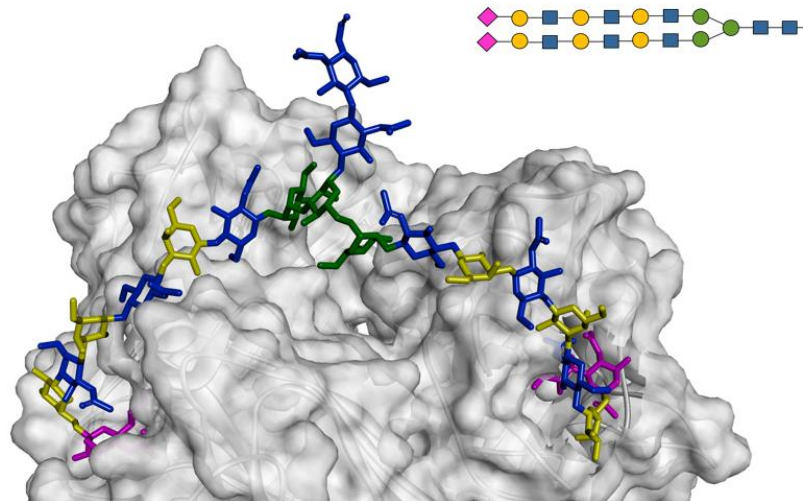


Figure S7. Model of the bidentate binding of a triLacNac glycan recognizing two sites of the same hemagglutinin trimer. A/CA/04/09 protein coordinates were obtained from the PDB 3UBE and were overlaid with the model of a triLacNac biantennary glycan in complex with H3N2 Victoria/11 built by Prof. Robert Woods group^[3] (coordinates kindly provided by Dr. Oliver Grant).

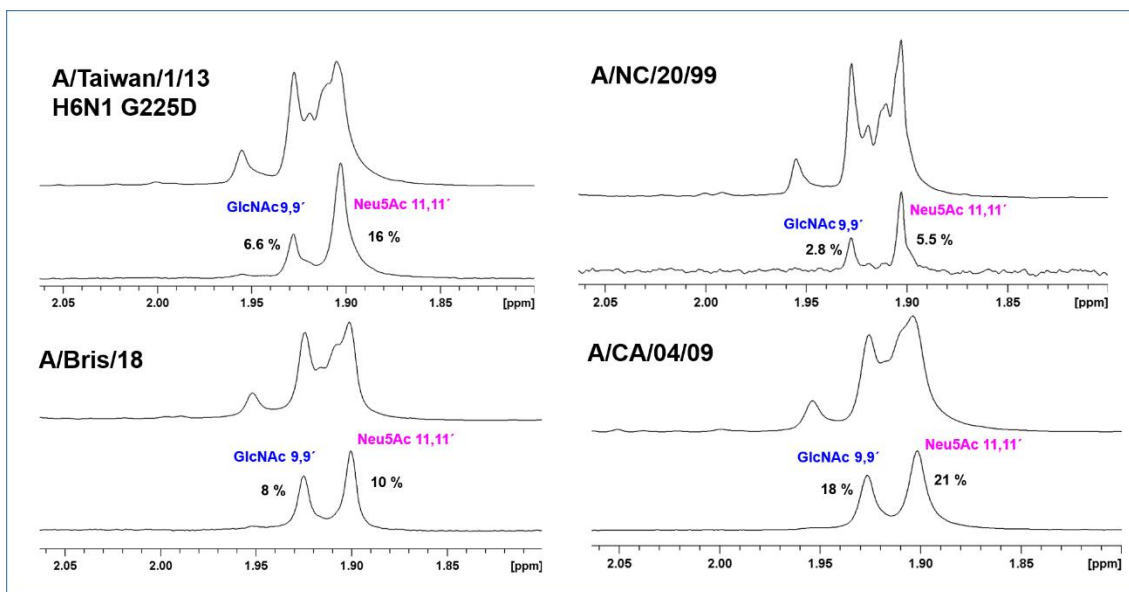


Figure S8. Comparison of the STD spectra (acetyl group region) of the TriLacNAc biantennary N-glycan in the presence of the four hemagglutinins characterized in this work. The effects are quantified comparing the intensity of each signal in the STD with the same signal in the off resonance spectrum.

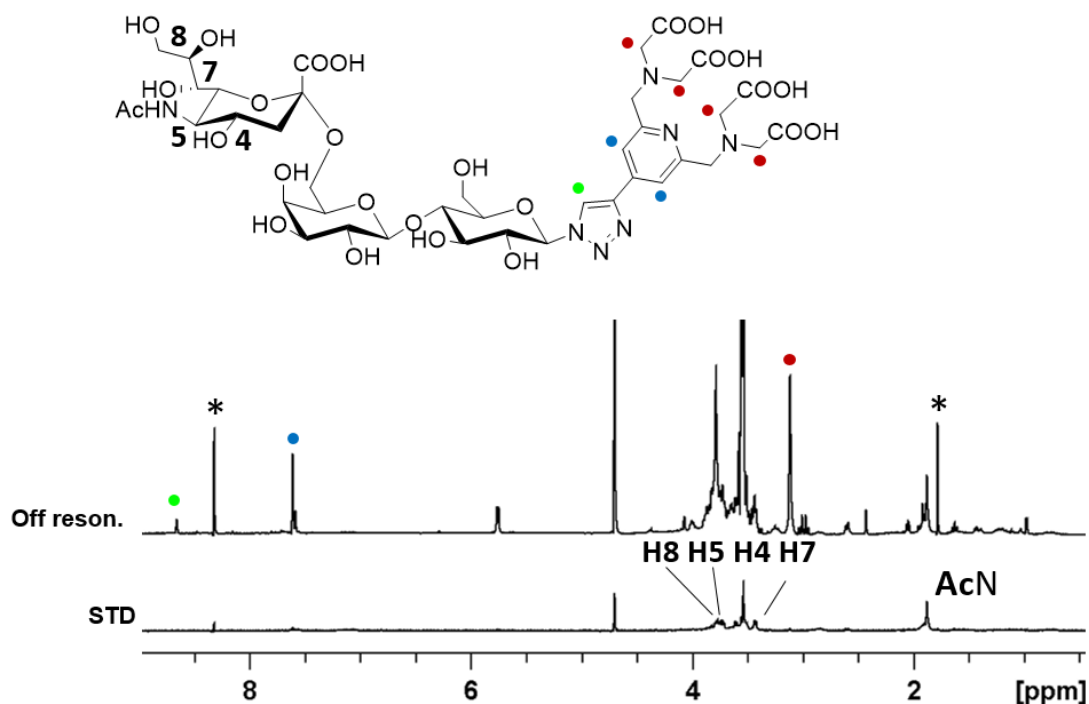


Figure S9. STD spectrum of the trisaccharide probe in the presence of H1 A/CA/04/09. The off resonance spectrum (reference experiment) is given for comparison. The signals of the chelating unit (labelled with circles) can not be detected in the STD experiment, pointing out that the tag does not interact with the protein. The signals of sialic acid (protons H4, H5, H7, H8 and AcN CH₃) are detected in STD as expected. Buffer impurities are labelled with asterisks.

- [1] M. Qi, M. Hulsmann, A. Godt, *Synthesis-Stuttgart* 2016, 48, 3773-3784.
- [2] A) T. Tanaka, Y. Zhou, C. Tamoto, Y. Kurebayashi, T. Takahashi, T. Suzuki, *J. Appl. Glycosci.* 1999, 64, 43-48. B) M. W. J. Whitehead, N. Khanzhin, L. Borsig, T. Hennet, *Cell Chem Biol* 2017, 24, 1336-1346 e1333.
- [3] W. Peng, R. P. de Vries, O. C. Grant, A. J. Thompson, R. McBride, B. Tsogtbaatar, P. S. Lee, N. Razi, I. A. Wilson, R. J. Woods, J. C. Paulson, *Cell Host Microbe* **2017**, 21, 23-34.